International Journal of Communication 17(2023), 2988–3011 1932–8036/20230005

Are People Easily Duped by Disinformation? Experimental Evidence for Open Vigilance

WUYAO DING
YAN GE[1]
Shanghai Jiao Tong University, China

Although it has been commonly believed that people are gullible, the open vigilance (OV) theory advocates that people in the default state of information processing are conservative. By deepening their processing, they likely look for true cues to accept information instead of false cues to reject it. The current paper examined whether the OV theory works in the evaluation of information credibility in two studies. Study 1 found that more elaborated information processing brought higher perceived information credibility, no matter whether the news was true or false. Study 2 found that false cues made little impact on perceived information credibility, but true cues increased it significantly. As for people who had little prior knowledge, enriching either true or false details increased their perceived credibility. These findings enhance the understanding of how people judge true and false news.

Keywords: disinformation, cognition, media literacy, judgment, and information processing

The widespread of online disinformation has been of much concern because it is evidently a threat to societal stability (Cosentino, 2020). To confront this danger, various strategies have been proposed (e.g., Lim, 2020; Pennycook, McPhetres, Zhang, Lu, & Rand, 2020). Most of those strategies are based on the credulity hypothesis, which assumes that people are gullible and vulnerable to information manipulation (Forgas & Baumeister, 2019; Shabi, Oyewusi, & Shabi, 2018). Therefore, elaborate thinking and skeptical reflection should be used to detect false information (Burkhardt, 2017; VanderBorght, 2009). Despite the popularity of such claims, the empirical supports are not as robust as expected (Chu, 2020; Jones-Jang, Mortensen, & Liu, 2019).

Wuyao Ding: dingwy18@sjtu.edu.cn
Yan Ge (corresponding author): yge@sjtu.edu.cn
Date submitted: 2021-12-31

[1] The authors are affiliated with the School of Media & Communication at Shanghai Jiao Tong University. This research was supported by a grant (21AXW006) from the National Social Science Foundation of China. The authors are grateful to Hugo Mercier for helpful explanations of the theory of open vigilance.

Recently, Mercier (2020) proposed the theory of OV to explain the cognitive process that distinguishes between true and false information. Quite different from the credulity hypothesis, the OV theory posits that people are conservative rather than credulous by default. With cognitive efforts increasing, they likely look for true cues to accept information instead of false cues to reject information. Since this interesting theory has been largely built with indirect evidence from the bottom up, this article is an effort to validate Mercier's (2020) principal assumptions directly with experiments, which focus on the effect of OV on content acceptance.

The Credulity Hypothesis

When saying "measure twice, cut once," Westerners are not alone. Chinese express the same idea by saying "think triple times before acting." More often than not, many believe that more elaborate thinking would improve decisions or judgments.

The dual-system theory (Evans, 2008; Kahneman, 2011) provides a formal model for this folk wisdom. It argues that humans employ two systems in processing information. System 1 works fast, unconsciously, in a modularized and heuristic manner, with low energy consumption but is prone to errors. The process of System 2 is slower, analytical, and reflective, which consumes more energy but produces fewer errors. Since humans are cognitive misers, they rely more on resource-saving System 1, which results in more misjudgments (Kahneman, 2011).[2]

Inspired by Spinoza (1677/1982), Gilbert (1991) linked the misjudgment caused by cognitive stinginess with gullibility, suggesting that information perception had two phases: an initial tendency to accept information, followed by efforts to certify or reject it. Because of cognitive stinginess, people more easily accept information without sufficient scrutiny and thus are credulous per se. Gilbert, Krull, and Malone (1990) and Gilbert, Tafarodi, and Malone (1993) also demonstrated the above assumption experimentally, and they observed that people were more likely to misjudge the falsehood as truth when they were under pressure from the cognitive load.

Limitations of the Credulity Hypothesis

Based on the credulity hypothesis, increasing cognitive effort to think elaborately has been believed by many as a way to improve people's ability to resist disinformation. A large number of media literacy theories and practical strategies have thus been recommended accordingly.

Does this work? Not necessarily. Empirically, Chu (2020) found that the more elaborately people thought, the more likely they were to fail in identifying false information. In an experiment on confession authenticity judgment, Kassin, Meissner, and Norwick (2005) discovered that police investigators with more relevant knowledge performed worse than college students. Schul, Mayo, Burnstein, and Yahalom (2007) showed that participants who attributed uncertainty to deception performed worse than those who attributed uncertainty to chance in the task of predicting outcomes because when participants

[2] System 1 and System 2 are also referred to as Type 1 and Type 2, Heuristic and Analytic, and other names by scholars (Hu & Hu, 2012).

believed the outcomes were caused by humans, they tried hard to control the occasions and eliminate every error. Such a more elaborate strategy, in turn, decreased the accuracy of their judgment instead of enabling a better judgment.

A number of studies also found that the acceptance of false information may be not because people are gullible, as Gilbert (1991) suggested. Hasson, Simmons, and Todorov (2005) showed that the participants accepted false information when, and only when, the information they encountered was uninformative. Street and Richardson (2015) demonstrated that when being forced to make a binary judgment (true or false), the participants displayed truth bias, while the very bias disappeared as the participants were allowed to express the uncertainty they had felt. Richter, Schroeder, and Wöhrmann (2009) further showed that participants with validity-relevant beliefs were able to reject false information fast and efficiently even though they were under cognitive load.

Although these findings cast doubt on the credulity hypothesis, they were not sufficient to form an alternative theory until the OV theory was proposed.

The OV Theory

The OV theory was proposed by Mercier (2020) after he had developed the epistemic vigilance theory, which advocated that vigilance is a key component of human communication (Sperber et al., 2010). According to the epistemic vigilance theory, people have to communicate with each other to obtain information about their surroundings. The adoption of others' opinions usually means that the sources of the opinions are perceived to be honest and capable of knowing the truth (i.e., we should trust the sources). However, since the inaccurate information may be disseminated either for deception or because of the sources' limited competency, trust would likely come along with epistemic vigilance. In other words, trust rarely stands alone and must coexist with vigilance.

The mechanism of this coexistence has been more or less revealed in empirical studies. In a trust game experiment conducted by Fetchenhauer and Dunnin (2009), the participants were asked to decide how much money they wanted to hand over to other anonymous participants, under the condition that the others might give the money back or keep it for themselves. When guessing how much money the others would return, the participants were rather vigilant and made pessimistic estimations. Nevertheless, the same participants behaved more open-mindedly and the money they gave to the others was significantly more than the amount they expected to be returned. This implies that vigilance and trust appeared together in the decision-making process.

Since unreliable persons may transmit credible information, while reliable persons may disseminate inaccurate information, epistemic vigilance also requires directly examining information content. People inspect encountered information based on their background beliefs, a process considered to be instinctive and effortless (Mercier, 2017). If the external information is not consistent with their preexisting beliefs, they either directly reject the information or apply cognitive efforts to decide whether their beliefs should be revised (Mercier, 2017, 2020; Sperber et al., 2010). People are, so to say, hesitant to accept new information to avoid deception.

There has been sufficient evidence that people are conservative with external information. For example, in a classical study on advice-taking, Yaniv (2004) asked participants to form initial opinions before receiving advice from others and analyzed the influence of external opinions on each participant's opinions. The study revealed that people assign more weight to their initial opinions than to those of others, and as the differences between the external opinions and their initial opinions increased, the weight assigned to external opinions decreased. It can be seen that it is often challenging to change one's opinion, and people are cautious to accept new information.

From an evolutionary perspective, Mercier (2020) suggested an analogy between the evolution of cognition for communication and the evolution of diet strategies. The creatures with lower intelligence, as he pictured it, tended to adopt a single and conservative dietary strategy because it could guarantee food safety. In the meantime, they might face the danger of extinction if the food supply they depended on sharply decreased. In contrast, creatures of higher intelligence likely adopted complex dietary strategies and had diverse food supplies to be more adaptive. Considering the likelihood of eating harmful foods, the omnivorous strategy required the agents to have the propensity to be open on the one hand and vigilant on the other hand. Mercier (2020) pointed out that similar to diet strategies, the cognitive mechanisms for communication evolved from being extremely conservative (i.e., accepting specific signals only) to being open but vigilant about a variety of signals. To highlight that cognitive openness does not translate to less vigilance, Mercier (2020) proposed the concept of OV as a replacement for epistemic vigilance.

In explaining the mechanism of OV, Mercier (2020) argued that humans are instinctively conservative and tend to reject information that is inconsistent with their preexisting beliefs, mostly to prevent being deceived by falsehood. Rejecting information in a changing world, however, can hinder adaptation and learning. Therefore, conditional openness is an effective supplement that enables people to consider adjusting their beliefs and accepting new information by identifying supportive cues.

Research Framework

The OV theory is an epistemological theory primarily based on psychology, linguistics, and philosophy, which explains the mechanisms of decision-making related to believing informants and information. The theory accounts for two critical assumptions: (1) people are instinctively conservative in the default state, not credulous, and this conservativeness is likely to soften as the process becomes more elaborate, and (2) people are more likely to seek true cues to accept information instead of seeking false cues to reject it.

Although there is some multidisciplinary evidence that supports the theory indirectly (Mercier, 2017, 2020), the present paper intended to empirically test the theory more directly, focusing on the effects of open vigilance on news content acceptance. To measure the effect, we used perceived information credibility.

This measurement needs more explanation. As demonstrated in the literature, truth rating (e.g., Nadarevic, Reber, Helmecke, & Kose, 2020), authenticity rating (e.g., Gaozhao, 2021), accuracy rating (e.g., Pennycook & Rand, 2020), and credibility rating (e.g., Mena, Barbe, & Chan-Olmsted, 2020) are extensively used to measure the degree of information acceptance (Bryanov & Vziatysheva, 2021).

However, the credibility of information, which is defined as the veracity judgment of the communication content, can encompass a set of associated concepts, such as believability, accuracy, and authenticity (Appelman & Sundar, 2016). We followed this convention.

Based on the assumption of the OV theory that a new message would be taken as a falsehood by default and that accepting it would need more effort, we proposed our first hypothesis for Study 1:

H1: *Compared with less effort in information processing, more effort increases perceived information credibility.*

Study 1 randomly divided participants into three groups: control, more-questions, and read-aloud. The control group evaluated the credibility of the news right after reading it. The more-questions group evaluated the credibility of the news after reading it and answering a few questions related to the news. The read-aloud group evaluated the information credibility of the news after reading the information aloud. It has been documented in the literature of cognitive studies that completing a more-questions task requires additional efforts in terms of thinking about the message than just reading it (Craik & Tulving, 1975) and that reading a message aloud stimulates additional effort for encoding the information (Reynolds & Besner, 2006), which made the information processed more than just reading it silently (Craik & Lockhart, 1972; Kelly, Ensor, Lu, MacLeod, & Risko, 2022; Leow, 2015). Therefore, comparing the perceived information credibility from the control group with that from the more-questions and read-aloud groups can validate H1.

Based on the second assumption of the OV theory (i.e., people are likely to look for true cues to accept information instead of false cues to reject it), we proposed a set of hypotheses for Study 2:

H2a: *The perceived credibility of information increases when true cues are received for the information.*

H2b: *The perceived credibility remains unchanged when false cues are received for the information.*

In Study 2, the participants were first asked to read a vague version of a message and rate its credibility; this provided the baseline ratings for the perceived credibility of the message. The participants were then randomly divided into two groups: (1) one group of participants read a new version of the message embedded with a true cue (i.e., true version), and (2) the other group of participants read a new version of the message embedded with a false cue (i.e., false version). After reading the assigned new versions of the message, the two groups rated the information credibility of the versions they just read. The validity of the hypotheses was determined by whether the participants who received the true cues increased their credibility ratings relative to their baseline credibility ratings, and whether the participants who received the false cues remained unchanged in their credibility ratings relative to their baseline credibility ratings.

Study 1

Study 1 aimed at testing whether more elaborate information processing would result in a higher rating of the perceived credibility of information (H1).

Method

Design

Designed as a one-way between-subjects experiment, Study 1 used the depth of information processing as the independent variable and the perceived information credibility as the dependent variable. The information processing was manipulated at three levels: Control group, more-questions group, and read-load group (Figure 1).

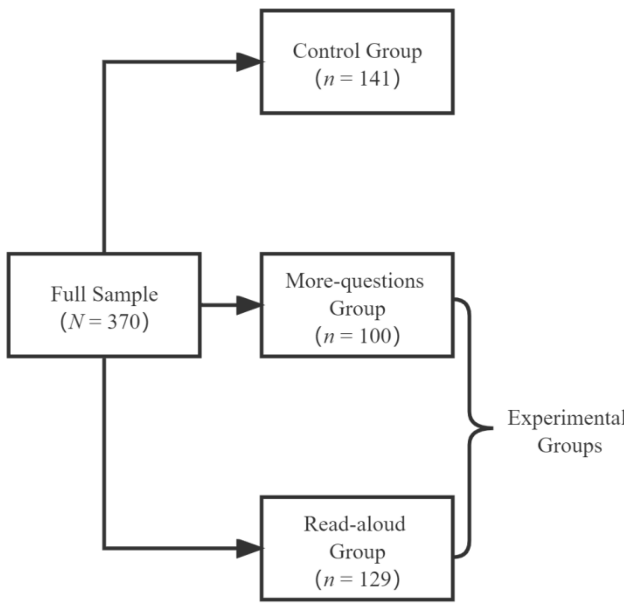

Figure 1. The grouping of participants in Study 1.

Although a few scales include several items to evaluate message credibility (e.g., Li & Suh, 2015), we used a simple and generic question (i.e., *How credible do you think the message is?*) similar to Appelman and Sundar's question ("How credible was the article you just read"; 2016, p. 68) to measure participants' responses. This allowed us to examine the participants' intuitive perception of credibility since answering more items may lead the participants to think more elaboratively about the information (Craik & Lockhart, 1972). The credibility was scored on a scale of 1 to 6 ("very incredible" to "very credible") to prevent the neutral option preference.

Stimuli

We adopted and adapted four real (false/true) news items that had been professionally verified as stimuli. Two of the news items were true: "Beijing Time Originates from Shaanxi Rather Than Beijing" (CCTV

News Center, 2021; abbreviated as *Beijing Time*) and "Taking Pictures with Hands' Gesture of 'V' Will Give Away Your Fingerprints" (Qnews, 2019; abbreviated as *Fingerprints*). The other two news items were false: "Russian Elementary School Students Return Home from School in a Severe Blizzard" (China Fact Check, 2020; abbreviated as *Blizzard*) and "Hundreds of Thousands of People Were Killed as a Result of Excessive PM2.5 Generated from Rural Cooking" (Han, 2021; abbreviated as *PM2.5*). All the news items were presented in headline, text, and image format to match the reading habits of media consumers (an example is provided in Appendix A, Figure A1).

The two false items were not completely baseless. In the case of *Blizzard*, false information comprised certain key objects mixed with incorrect details, such as the wrong location name for the news event (China Fact Check, 2020). The story of *PM2.5* was based on a scientific investigation that reported on the association between rural residential emissions and premature deaths in China (Yun et al., 2020). However, the false version of the news distorted the scientific investigation by misrepresenting the harmful effects of solid fuel emissions as being caused by rural cooking, and falsely equating premature deaths with deaths (Han, 2021).

Participants

A total of 378 Chinese nationals[3] were recruited via Credamo (an online platform widely adopted by institutions for conducting research in China) during August 17–20, 2021. Among them, 370 passed the quality checks (see the subsection Quality Check for details) and participated in the experiment. Of those participants, 202 were male, and 168 were female. One had a junior secondary diploma, 18 a senior secondary diploma, 43 a junior college diploma, 282 a bachelor's degree, 21 a master's degree, and five a doctorate. They were classified into five age groups: 20 in the less-than-20 group, 242 in the 21–30 group, 92 in the 31–40 group, 13 in the 41–50 group, two in the 51–60 group, and one in the more-than 60 group.

Procedure

After reading the informed consent and the procedure introduction, the participants were randomly assigned to the control group and two experimental groups. They then read the four news items in random order.

The control group evaluated the credibility of each news item immediately after reading it. The more-questions group, after reading each news item, was asked to answer several additional questions

[3] We estimated the sample size before sampling with a desired power of 0.95, an effect size of 0.25, and an alpha error of 0.05 (Tool Use and Background Knowledge as the control variable, the manipulation method as the independent variable, and the perceived information credibility as the dependent variable). The calculations indicated a sample size of 251. Considering the uncertainty in online experiments, we decided to recruit 1.5 times more participants. The method for determining sample size in Study 2 was similar to this approach, but due to the increased difficulty in completing the experiments in Study 2 compared with Study 1, we recruited two times more participants than estimated. The actual grouping results for both two studies are provided in Appendix B.

related to the information (such as "How important do you think this piece of information is?") and then they evaluated the item's credibility. The read-aloud group was asked to read aloud the news item (audios of the participants reading aloud were recorded) before the evaluation.

After they evaluated all the items, the participants were asked to answer whether they had used tools such as search engines in making judgments (Tool Use) and whether they had had prior knowledge relevant to the items they had just read (Background Knowledge). At the end of the questionnaire, they were asked to report their demographic information.

As the experiment was completed, the purpose of the study and the manipulation of the true and false news were revealed. The participants were thanked for their participation.

Quality Check

We used two methods for quality checks. First, the participants were given five options under Background Knowledge: "Have you previously learned of any of the above information? If yes, please select that information, and if no, select 'none of the above.'" Of the options, four showed the titles of the news items, and the fifth was "none of the above." If participants selected both a news title option and "none of the above," they failed the check since they did not answer the questions with sufficient attentiveness. One in the control group, one in the more-questions group, and one in the read-aloud group were excluded by this check. Second, for the read-aloud group, the audio recordings of the participants were screened to check whether they had read every item loudly. Five were excluded by this check. As a result, a total of eight people were excluded by the quality checks.

Since the data excluded are less than 3% (i.e., eight dropouts of 378), the attrition can be reasonably regarded as random, and this dropout rate is thus acceptable (Nunan, Aronson, & Bankhead, 2018). The dropout analysis also shows that there were no systematic differences in the number of valid and dropout participants among the three groups ($x^2(2) = 5.517$, $p = .063$).

Results

Preliminary Analysis

Descriptive statistics are shown in Table 1.

Table 1. Descriptive Statistics of Study 1.

	Perceived Information Credibility ($M \pm SD$)		
News Item	Control Group (n = 141)	More-Questions Group (n = 100)	Read-Aloud Group (n = 129)
Beijing Time	4.18 ± 1.64	4.98 ± 1.08	4.81 ± 1.19
Fingerprints	3.78 ± 1.60	4.55 ± 1.40	4.46 ± 1.32
Blizzard	3.65 ± 1.56	4.08 ± 1.58	4.25 ± 1.47
PM2.5	2.64 ± 1.61	3.71 ± 1.66	3.3 ± 1.69
All the items	14.26 ± 3.79	17.32 ± 3.85	16.81 ± 3.33

We conducted a preliminary analysis with analysis of variance (ANOVA) on the data from the participants who passed the quality check. The manipulation of information processing was used as the independent variable and the perceived information credibility (the sum of the four items) as the dependent variable.

The results showed a significant effect of the independent variable on the dependent viable [F (2, 367) = 25.82; $p < .001$; partial $\eta^2 = .123$]. More specifically, the scores given by both the more-questions group (M = 17.32, SD = 3.85) and the read-aloud group (M = 16.81, SD = 3.33) were significantly higher than the score by the control group (M = 14.26, SD = 3.79), 95% confidence interval (CI) for the difference mean between the more-questions group and the control group was [2.13, 4.0 0], and 95% CI for the difference mean between the read-aloud group and the control group was [1.68, 3.43]. There was no significant difference between the read-aloud and the more-questions groups. This allowed us to combine data produced by the two groups into a single group, namely, the experimental group, for further analysis.

Controlled Analysis

To avoid the possible influence of Tool Use and Background Knowledge on the perceived information credibility (Chu, 2020), we controlled these variables and performed ANOVA again for each of the news items. The results (Table 2) showed a significant effect of the manipulation of the information process on the items of *Beijing Time* (true) [F (1, 362) = 16.17, $p < .001$, partial $\eta^2 = .043$], *Fingerprints* (true) [F (1, 362) = 4.05, $p = .045$, partial $\eta^2 = .011$], and *PM2.5* (false) [F (1, 362) = 11.70, $p = .001$, partial $\eta^2 = .031$], and the scores given by the participants in the experimental group were significantly higher than the score by the control group.

Nonetheless, the manipulation did not produce a significant effect on *Blizzard* (false) [F (1, 362) = 3.24, $p = .073$, partial $\eta^2 = .009$] at the significance level of 0.05 although the perceived credibility increased significantly at the level of 0.1. We speculate that the item might have been presented in a manner that looked so true that people preferred to take it as credible. However, considering the item is false actually, the possibility of space for enhancing its credibility is finite (for a falsehood, there would be limited supportive cues to be found even if enough elaborate effort was put in). That is why the experiment effect of this news item was not as strong as that of the other three items.

Table 2. The Manipulation Effect on the Perceived Information Credibility.

	Perceived Information Credibility ($M \pm SD$)		
News Item	Control Group (n = 141)	Experimental Group (n = 229）	p Value
Beijing Time	4.18 ± 1.64	4.88 ± 1.15	< .001***
Fingerprints	3.78 ± 1.60	4.50 ± 1.35	.045*
Blizzard	3.65 ± 1.56	4.17 ± 1.52	.073
PM2.5	2.64 ± 1.61	3.48 ± 1.69	.001**

Note. ***$p < .001$; **$p < .01$; *$p < .05$.

These findings suggest that more elaborate information processing is correlated with a higher perception of information credibility, which largely supports H1.

Study 2

Using H2a and H2b, Study 2 aimed at examining whether true/false cues would influence perceived information credibility.

Method

Design

Study 2 had a mixed-design experiment, with treatment time (before and after the treatment) and cue type (true and false) as the independent variables, and perceived information credibility as the dependent variable. The treatment time was a within-subjects variable, and every participant was subjected to it. The cue type was a between-subjects variable, and the participants were subjected to the manipulations of either a true or a false cue.

The credibility was measured with the simple and generic question (i.e., *How credible do you think the message is?*) with scoring from 1 to 6 ("very incredible" to "very credible"), as in Study 1. The cue was manipulated as the details that influence people's perception of news credibility.

To make the experimental results robust, we performed experiments with three different sets of materials by the same experimental design. The participants were randomly divided into three groups with different news items. Each group was asked to read and evaluate the credibility of one item assigned to them. Then, each of the groups was further divided into two subgroups: the true cue subgroup and the false cue subgroup. The true cue subgroup was treated with true cues relevant to the news item they just had read, and the false cue subgroup was treated with false cues. The participants in both subgroups were asked to read those cues, respectively, and then evaluate the information's credibility again (Figure 2).

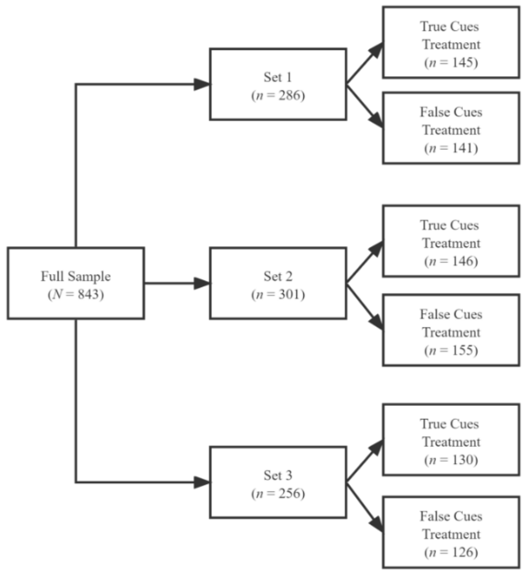

Figure 2. The grouping of participants in Study 2.

Stimuli

We created three sets of materials, namely Set 1, Set 2, and Set 3, each containing three versions of a news item: vague, true, and false. The three versions had the same structure, with only a few words changed to manipulate the validity of the stories. An example is shown in Appendix A, Figure A2.

Of Set 1, the three versions of the item were created based on the article "Your Blood Type Can Change?!" (DXY, 2016). The only difference among them was the cause of change in blood types. The vague version described the cause as *a specific medical intervention*; the true version, as *a hematopoietic stem cell transplant*; and the false version, as *a heart transplant*.

Of Set 2, the three versions of the item were created based on the article "Diamonds Can Be Made from Hair" (Tadpole Stave, 2020). These versions differed only in the element of hair used to make diamonds. The vague version described it as *hair has the same element as diamond*, the true version as *hair has the same carbon element as diamond*, and the false version as *hair has the same nitrogen element as diamond*.

Of Set 3, the three versions of the item were created based on the report "Russian Elementary School Students Insist on Attending Classes at Minus 51℃" (CCTV.com, 2020). They were different only with regard to the location of the event. The vague version described the location as *the local area*, the true version as *Oymyakon village in eastern Siberia of Russia*, and the false version as *Brody Town in Lille of France* (a fictional place).

Participants

During August 22–25, 2021, a total of 1,036 Chinese nationals were recruited via Credamo. Of these, 843 passed the quality checks (details provided in the subsection Quality Check) and participated in the experiment. Among the participants, 373 were male, and 470 were female. Regarding education level, 51 had a senior secondary diploma, 111 had a junior college diploma, 596 had a bachelor's degree, 70 had a master's degree, and nine had a doctorate, while the rest had a junior secondary school diploma or below. According to age, 65 were less than 20 years of age, 499 were aged 21–30, 227 were aged 31–40, 38 were aged 41–50, and 14 were aged 51–60. They were randomly assigned to three sets, as shown in Figure 2.

Procedure

The procedure was displayed on five screens. As the participants entered Study 2 via either personal computer or smartphone, screen 1 asked them to read the informed consent form and the procedure instructions.

On screen 2, all the participants read the vague version of the news items assigned to them and rated their credibility.

On screen 3, they were provided with either the true or false versions of the news item they had just read. To ensure that they read the material carefully, they were asked whether they had noticed differences between the items on the current and the previous screens. They then rated the credibility of the news item on the current version.

On screen 4, they were asked to provide their demographic information and to answer whether they had used tools such as search engines to assist in making their judgments (Tool Use), and whether they had had prior knowledge about the news items they just read (Background Knowledge).

On screen 5, the participants were informed of the purpose of the study, and the manipulations. Then they were appreciated for their participation.

Quality Check

Two methods were used for quality checks. One was performed by asking the participants if they had noticed the differences between the vague version and the other version. The participants passed the quality check if they were able to correctly remember and locate the differences. There were 171 pieces of data excluded for this check: 65 were in Set 1, 61 in Set 2, and 45 in Set 3.

In addition, the consistency of the responses to Tool Use and Background Knowledge was used to test the robustness of the responses. The participants were asked "Did you use tools (e.g., search engines) to assist you in making your judgment during this response? Or have you ever heard anything about it?" They were given four options: "With tools," "Without tools," "Have heard this information," and "Have never heard this information." If anyone selected both "with tools" and "without tools," or both "heard about this information" and "never heard about this information," their responses were excluded from the later analysis. There were 22 pieces of data excluded for this check: 11 were in Set 1, eight in Set 2, and three in Set 3.

As a result, there were 76 (in Set 1), 69 (in Set 2), and 48 (in Set 3) pieces of data that failed the checks. The dropout analysis shows that there were no systematic differences in the number of valid and dropout participants between the manipulation groups ($x^2(5) = 5.036$, $p = .411$).[4]

Results

We conducted a preliminary analysis to detect the effects of the treatment on the perceived information credibility and a control analysis to eliminate the possible effect of Tool Use and Background Knowledge on the dependent variable. Descriptive statistics results are shown in Table 3.

Table 3. Descriptive Statistics of Study 2.

				Treatment	
	Set	Cue	*n*	Before	After
Preliminary analysis	1	T	145	2.83 ± 1.59	3.92 ± 1.60
		F	141	2.65 ± 1.51	2.69 ± 1.66
	2	T	146	2.77 ± 1.50	3.50 ± 1.56
		F	155	2.60 ± 1.32	2.55 ± 1.42
	3	T	130	2.99 ± 1.52	4.22 ± 1.34
		F	126	2.98 ± 1.60	3.04 ± 1.70
Control analysis	1	T	108	2.47 ± 1.45	3.55 ± 1.60
		F	103	2.46 ± 1.30	2.79 ± 1.58
	2	T	105	2.33 ± 1.18	3.06 ± 1.45
		F	110	2.42 ± 1.12	2.65 ± 1.36
	3	T	99	2.70 ± 1.42	4.05 ± 1.30
		F	96	2.68 ± 1.43	3.03 ± 1.63

Note. T= true, F = false.

Preliminary Analysis

Repeated-measures ANOVA was conducted with the treatment time as a within-subjects variable, the cue type as a between-subjects variable, and the perceived information credibility as the dependent variable for all the three sets of materials (see Figure 3a).

[4] The proportion of excluded participants was higher in Study 2 than in Study 1. This discrepancy is likely due to the fact that Study 2 was more demanding in terms of cognitive resources, which may have resulted in a higher likelihood of participants making errors and thus being excluded. Nevertheless, the dropout rate is acceptable since there were no significant differences between valid and dropout participants.

For Set 1, the effects of both the treatment time [F (1, 284) = 33.73, $p < .001$, partial $\eta^2 = .106$] and the cue type [F (1, 284) = 19.25, $p < .001$, partial $\eta^2 = .063$] were observed. The interactive effect of the treatment and the cue type [F (1, 284) = 28.85, $p < .001$, partial $\eta^2 = .092$] was also observed.

More specifically, before the cues were added, no significant difference in the perceived credibility was found between the true- and the false-cue groups [F (1, 284) = .99, $p = .321$, partial $\eta^2 = .003$]. The difference emerged when the true and false cues were added to the two groups, respectively [F (1, 284) = 40.73, $p < .001$, partial $\eta^2 = .125$]. Moreover, for the true-cue group, the perceived credibility was significantly higher after receiving the true cues [F (1, 284) = 63.37, $p < .001$, partial $\eta^2 = .182$]. For the false-cue group, no such change was observed [F (1, 284) = .09, $p = .759$, partial $\eta^2 < .001$].

For Set 2, the effects of both the treatment time [F (1, 299) = 18.55, $p < .001$, partial $\eta^2 = .058$] and the cue type [F (1, 299) = 14.47, $p < .001$, partial $\eta^2 = .046$] on the perceived credibility were found. The interactive effect of the treatment time and the cues type on the perception was also significant [F (1, 299) = 23.80, $p < .001$, partial $\eta^2 = .074$].

Moreover, before the cues were added, no significant difference in the perceived credibility was found between the true- and false-cue groups [F (1, 299) = 1.15, $p = .285$, partial $\eta^2 = .004$]. The difference emerged between the two groups after they had received true and false cues, respectively [F (1, 299) = 30.39, $p < .001$, partial $\eta^2 = .092$]. Compared with its credibility perception score before the true cues were added, the true-cue group produced a significantly higher score of the credibility perception [F (1, 299) = 40.96, $p < .001$, partial $\eta^2 = .120$] after the true cues were added. This was not the case for the group that received false cues [F (1, 299) = .17, $p = .682$, partial $\eta^2 < .001$].

For Set 3, the effects of both the treatment time [F (1, 254) = 37.37, $p < .001$, partial $\eta^2 = .128$] and the cue type [F (1, 254) = 13.84, $p < .001$, partial $\eta^2 = .052$] on the perceived credibility were observed. The interactive effect of the treatment time and the cue type was also significant [F (1, 254) = 30.39, $p < .001$, partial $\eta^2 = .107$].

Moreover, before the participants were treated with the true and false cues, there was no significant difference between the true-cue and the false-cue groups [F (1, 254) = .01, $p = .934$, partial $\eta^2 < .001$]. After receiving the true and false cues, the difference appeared [F (1, 254) = 38.49, $p < .001$, partial $\eta^2 = .132$]. Compared with the initial credibility score, a higher score was found from the true-cue group [F (1, 254) = 68.65, $p < .001$, partial $\eta^2 = .213$] and no significant score changed from the false-cue group [(F (1, 254) = .18, $p = .674$, partial $\eta^2 = .001$].

These findings show that people perceived the information as more credible when they noticed the true cue relevant to the information, but the false cue could not produce the same effect. H2a and H2b are therefore supported, respectively.

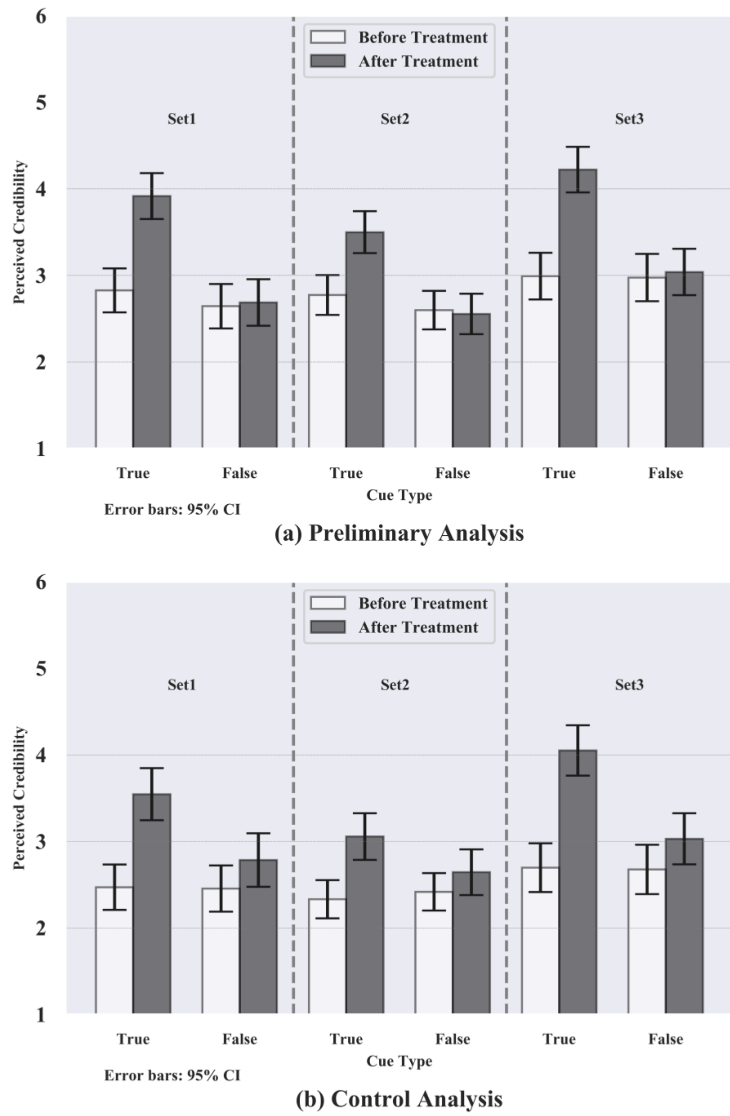

Figure 3. The cue treatment effects on perceived information credibility.

Control Analysis

To avoid the possible influence of Tool Use and Background Knowledge on the perceived credibility, repeated-measures ANOVA was performed only on the participants who had not used a search tool and had no background knowledge. Again, the treatment time was the within-subjects variable, the cue type was the between-subjects variable, and the perception of information credibility was the dependent variable (see Figure 3b).

For Set 1, the effects of both treatment time [F (1, 209) = 41.93, p < .001, partial η^2 = .167] and cue type [F (1, 209) = 4.98, p = .027 < .05, partial η^2 = .023] were significant. The interaction effect of the treatment time and the cue type was also significant [F (1, 209) = 11.77, p = .001 < .01, partial η^2 = .053].

Analysis revealed no significant difference between the true- and the false-cue groups before they were given the cues [F (1, 209) = .01, p = .933, partial η^2 < .001]. But the difference appeared between the two groups after they received the true and false cues, respectively [F (1, 209) = 12.05, p = .001, partial η^2 = .055]. However, unlike in the preliminary analysis, the perceived credibility in this control analysis became significantly higher after adding either the true cue [F (1, 209) = 50.25, p < .001, partial η^2 = .194] or the false cue [F (1, 209) = 4.53, p = .035< .05, partial η^2 = .021].

For Set 2, the effect of the treatment time on the perceived credibility was significant [F (1, 213) = 32.11, p < .001, partial η^2 = .131]. In contrast, the effect of the cue type on the perceived credibility was not significant [F (1, 213) = 1.13, p = .288, partial η^2 = .005] although the interaction effect of the treatment time and the cue type was also significant [F (1, 213) = 8.75, p = .003 < .01, partial η^2 = .039].

A simple-effect analysis further revealed no significant difference between the true- and the false-cue groups before those cues were added [F (1, 213) = .29, p = .589, partial η^2 = .001]; but after those cues were added, a significant difference appeared between the two groups [F (1, 213) = 4.63, p = .033 < .05, partial η^2 = .021]. Notably, the true-cue group gave a significantly higher score [F (1, 213) = 36.35, p < .001, partial η^2 = .146], and the false-cue group gave a marginally significant higher score [F (1, 213) = 3.76, p = .054, partial η^2 = .017].

For Set 3, the effects of both the treatment time [F (1, 193) = 56.84, p < .001, partial η^2 = .228] and the cue type [F (1, 193) = 8.91, p = .003 < .01, partial η^2 = .044] were significant. The interaction effect of the treatment time and the cue type [F (1, 193) = 19.47, p < .001, partial η^2 = .092] was significant as well.

A simple-effect analysis further revealed that no significant difference in the perceived credibility appeared between the true- and the false-cue groups before the cues were added [F (1, 193) = .01, p = .922, partial η^2 < .001], but the treatment produced a significant difference between the two groups [F (1, 193) = 23.29, p < .001, partial η^2 = .108]. The scores increased in either the group receiving the true cues [F (1, 193) = 72.54, p < .001, partial η^2 = .273] or the group receiving the false cues [F (1, 193) = 4.82, p = .029 < .05, partial η^2 =.024] were all significant.

This result indicates that the false cue, like the true cue, was able to significantly increase the score of the perceived credibility if the participants had no knowledge pertaining to the news items and did not use the search tool. H2b, which claimed false cues made no impact on perceived information credibility, is supported only conditionally.

Discussion and Concluding Remarks

In this paper, we demonstrated first that the participants were conservative in the default state of information processing, as shown by the lower credibility scores they gave to the information. But the participants became more open when the processing was more elaborate, as evidenced by the higher credibility scores they gave to the information. This finding supports one of the two critical assumptions of the OV theory: People are conservative in the default state of information processing as compared with a later stage of more elaborate processing.

Second, we showed that participants who noticed the true cues increased the scores of their perceived information credibility, while those who noticed the false cues, made no change in the scores of their perceived information credibility. These findings support another critical assumption of the OV theory: People look for true cues to accept information instead of looking for false cues to reject it.

Interestingly, we also found that the participants, without relevant knowledge about the given information or using tools in the test, increased the scores of their perceived information credibility as long as the additional cues appeared, whether the cues were true or false. This effect is neither discussed by Mercier (2020) nor observed in relevant studies of OV. We can only speculate on the mechanism behind the judgments of those participants. According to the OV theory, after the conservativeness in the default state, people likely look for true cues to accept given information. This consequently brings about the question of how they judge which cue is true and which is not. In Study 2, it was clear that both true and false cues increased the score of information credibility but only among the participants who neither had little knowledge about the given information nor used search engines in the test. They were, so to say, unable to know which cues were true or false. Under such conditions, a conceivable explanation is that those participants, as predicted in the OV theory, began to look for cues to adjust their initial decision. Because of participants' lack of relevant knowledge to judge the given information, cues, as long as they could enrich the details of the information, might generate a sense of illusion of the validity of the information in the minds of those participants, regardless of whether the cues were true or false. This is an arresting speculation for it provides a mechanism-driven interpretation of why and how adding false details to false news would likely raise the credibility of the news. This speculation is worthy of further investigation.

In sum, the findings of this paper largely support the propositions of the OV theory and compel reflection on the strategies advocated by most media literacy programs since influenced by the credulity hypothesis; the general doctrine of those strategies encourages people to take caution when encountering new information, and "be skeptical to recognize falsehood" before accepting information. Although this strategy may increase the possibility of rejecting false information, it meanwhile increases the chance of rejecting true information, especially in the default state where people are likely more conservative.

It must be acknowledged that this investigation has several limitations. First, we did not bring information attributes into consideration, which previous studies suggested could affect information judgment (e.g., Hilbig, 2012). Second, although our cohort was diverse, since it was drawn from a convenience sample of non-probability sampling, it skewed toward younger and more educated both in Study 1 and in Study 2, and slightly skewed toward males in Study 1 and females in Study 2. This is not representative of the demography of the Chinese population (National Bureau of Statistics of China, 2021) and limits the external validity of our studies. Future studies should address all these limitations of the current paper.

References

Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly, 93*(1), 59–79. doi:10.1177/1077699015606057

Bryanov, K., & Vziatysheva, V. (2021). Determinants of individuals' belief in fake news: A scoping review determinants of belief in fake news. *Plos One, 16*(6), 1–25. doi:10.1371/journal.pone.0253717

Burkhardt, J. M. (2017). Can we save ourselves? *Library Technology Reports, 53*(8), 22–28. Retrieved from https://www.journals.ala.org/index.php/ltr/article/view/6500/8639

CCTV News Center. (2021, January 4). *Shén me? Běijīng shíjiān bù zài Běijīng fā bō!* [What? Beijing time does not originate from Beijing!]. Retrieved from https://view.inews.qq.com/a/20210104A06NRK00

CCTV.com. (2020, December 15). *Bù kuì shì "zhàndòu mínzú!" É xiǎoxuéshēng líng xià 51℃ jiānchí shàngxué* [Russian elementary school students insist on attending classes at minus 51℃]. Retrieved from https://news.cctv.com/2020/12/15/ARTIN5KFPXjlmSfCEiDTyIsK201215.shtml

China Fact Check. (2020, December 30). *Éluósī xiǎoxuéshēng líng xià 50 dù dǐng bàofēngxuě huí jiā?* [Russian elementary school students return home from school in a severe blizzard. Is it true?]. Retrieved from https://chinafactcheck.com/?p=2768

Chu, Y. (2020). Rénmen wèihé xiāngxìn bùshí xìnxī: Kēxué chuánbō shìjiǎo xià de rènzhī piānchà yǔ xìnxī jiànbiélì yánjiū [Why do we fall for misinformation? Cognitive bias and misinformation discrimination from the perspective of science communication]. *Xīn Wén Dà Xué, 175*(11), 71–87, 132. Retrieved from http://www.cnki.com.cn/Article/CJFDTotal-XWDX202011008.htm

Cosentino, G. (2020). *Social media and the post-truth world order: The global dynamics of disinformation*. Cham, Switzerland: Palgrave Macmillan. doi:10.1007/978-3-030-43005-4

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684. doi:10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268–294. doi:10.1037/0096-3445.104.3.268

DXY. (2016, August 27). *Xuèxíng jū rán huì biàn?! Guān yú xuèxíng de 5 gè zhēn jiǎ chuán yán* [Your blood type can change?! Five true and false myths about blood types]. Retrieved from https://dxy.com/article/7147

Evans, J. S. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*(1), 255–278. doi:10.1146/annurev.psych.59.103006.093629

Fetchenhauer, D., & Dunning, D. (2009). Do people trust too much or too little? *Journal of Economic Psychology, 30*(3), 263–276. doi:10.1016/j.joep.2008.04.006

Forgas, J. P., & Baumeister, R. (2019). *The social psychology of gullibility: Conspiracy theories, fake news and irrational beliefs*. New York, NY: Routledge. doi:10.4324/9780429203787

Gaozhao, D. (2021). Flagging fake news on social media: An experimental study of media consumers' identification of fake news. *Government Information Quarterly, 38*(3), 1–13. doi:10.1016/j.giq.2021.101591

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*(2), 107–119. doi:10.1037/0003-066X.46.2.107

Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology, 59*(4), 601–613. doi:10.1037/0022-3514.59.4.601

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology, 65*(2), 221–233. doi:10.1037/0022-3514.65.2.221

Han, F. (2021, January 8). *Nóngcūn shāo fàn chǎnshēng de PM2.5 dǎo zhì měi nián shù shí wàn rén sǐ wáng, shì zhēnde ma?* [Hundreds of thousands of people were killed as a result of excessive PM2.5 generated from rural cooking. Is this true?]. Retrieved from https://ppfocus.com/0/scf28f584.html

Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe it or not: On the possibility of suspending belief. *Psychological Science, 16*(7), 566–571. doi:10.1111/j.0956-7976.2005.01576.x

Hilbig, B. E. (2012). How framing statistical statements affects subjective veracity: Validation and application of a multinomial model for judgments of truth. *Cognition, 125*(1), 37–48. doi:10.1016/j.cognition.2012.06.009

Hu, Z., & Hu, X. (2012). Evans shuāngchóng jiāgōng lǐlùn de fāzhǎn guòchéng jiǎnyào shùpíng [A review of Evans's dual process theory]. *Xīn Lǐ Xué Tàn Xīn, 32*(4), 310–316. Retrieved from http://en.cnki.com.cn/Article_en/CJFDTOTAL-XLXT201204003.htm

Jones-Jang, S. M., Mortensen, T., & Liu, J. (2019). Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. *American Behavioral Scientist, 65*(2), 371–388. doi:10.1177/0002764219869406

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kassin, S. M., Meissner, C. A., & Norwick, R. J. (2005). "I'd know a false confession if I saw one": A comparative study of college students and police investigators. *Law and Human Behavior, 29*(2), 211–227. doi:10.1007/s10979-005-2416-9

Kelly, M. O., Ensor, T. M., Lu, X., MacLeod, C. M., & Risko, E. F. (2022). Reducing retrieval time modulates the production effect: Empirical evidence and computational accounts. *Journal of Memory and Language, 123*(1), 1–14. doi:10.1016/j.jml.2021.104299

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. New York, NY: Routledge. doi:10.4324/9781315887074

Li, R., & Suh, A. (2015). Factors influencing information credibility on social media platforms: Evidence from Facebook pages. *Procedia Computer Science, 72*(1), 314–328. doi:10.1016/j.procs.2015.12.146

Lim, S. (2020). Academic library guides for tackling fake news: A content analysis. *The Journal of Academic Librarianship, 46*(5), 1–11. doi:10.1016/j.acalib.2020.102195

Mena, P., Barbe, D., & Chan-Olmsted, S. (2020). Misinformation on Instagram: The impact of trusted endorsements on message credibility. *Social Media + Society, 6*(2), 1–9. doi:10.1177/2056305120935102

Mercier, H. (2017). How gullible are we? A review of the evidence from psychology and social science. *Review of General Psychology, 21*(2), 103–122. doi:10.1037/gpr0000111

Mercier, H. (2020). *Not born yesterday: The science of who we trust and what we believe*. Princeton, NJ: Princeton University Press. doi:10.1515/9780691198842

Nadarevic, L., Reber, R., Helmecke, A. J., & Kose, D. (2020). Perceived truth of statements and simulated social media postings: An experimental investigation of source credibility, repeated exposure, and presentation format. *Cognitive Research: Principles and Implications, 5*(1), 1–16. doi:10.1186/s41235-020-00251-4

National Bureau of Statistics of China. (2021). *China statistical yearbook-2021*. Retrieved from http://www.stats.gov.cn/tjsj/ndsj/

Nunan, D., Aronson, J., & Bankhead, C. (2018). Catalogue of bias: Attrition bias. *BMJ Evidence-Based Medicine, 23*(1), 21–22. doi:10.1136/ebmed-2017-110883

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science, 31*(7), 770–780. doi:10.1177/0956797620939054

Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality, 88*(2), 185–200. doi:10.1111/jopy.12476

Qnews. (2019, September 16). *Pāizhào bǐ jiǎndāo shǒu huì xièlòu zhǐwèn xìnxī? Zhuānjiā chēng rìcháng pāizhào yǐngxiǎng bù dà* [Can taking pictures with a hand gesture of "V" give away your fingerprints? An expert said it's possible, but the daily impact is not significant]. Retrieved from https://view.inews.qq.com/a/20190916A0I72800

Reynolds, M., & Besner, D. (2006). Reading aloud is not automatic: Processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance, 32*(6), 1303–1323. doi:10.1037/0096-1523.32.6.1303

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology, 96*(3), 538–558. doi:10.1037/a0014038

Schul, Y., Mayo, R., Burnstein, E., & Yahalom, N. (2007). How people cope with uncertainty due to chance or deception. *Journal of Experimental Social Psychology, 43*(1), 91–103. doi:10.1016/j.jesp.2006.02.015

Shabi, I. N., Oyewusi, F., & Shabi, O. (2018). Pattern of media literacy skills and Internet health information utilization among in-school adolescents in Osun State, south-west, Nigeria. *Nigerian Journal of Family Practice, 9*(1), 29–38. Retrieved from https://www.ajol.info/index.php/njfp/article/view/166270

Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359–393. doi:10.1111/j.1468-0017.2010.01394.x

Spinoza, B. (1982). *The ethics and selected letters* (S. Shirley, Trans.). Indianapolis, IN: Hackett Publishing Company. (Original work published in 1677)

Street, C. N. H., & Richardson, D. C. (2015). Descartes versus Spinoza: Truth, uncertainty, and bias. *Social Cognition, 33*(3), 227–239. doi:10.1521/soco.2015.33.2.2

Tadpole Stave. (2020, December 22). *Nǐ zhīdào ma, tóufà kě yǐ biàn zuànshí!!!* [Diamonds can be made from hair]. Retrieved from https://www.kedo.net.cn/c/1822/1822904.shtml

VanderBorght, M. A. (2009). *Teaching young children to be sophisticated media consumers*. University of Virginia. Retrieved from https://libraetd.lib.virginia.edu/public_view/cv43nx16p

Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes, 93*(1), 1–13. doi:10.1016/j.obhdp.2003.08.002

Yun, X., Shen, G., Shen, H., Meng, W., Chen, Y., Xu, H., . . . Tao, S. (2020). Residential solid fuel emissions contribute significantly to air pollution and associated health impacts in China. *Science Advances, 6*(44), 1–7. doi:10.1126/sciadv.aba7621

Appendix A: Examples of Experimental Stimuli*

Beijing Time Originates from Shaanxi rather than Beijing

Beijing time is a phrase commonly seen in our daily lives. However, Beijing time is actually a definition that refers to UTC/GMT+08:00, a time zone for the capital of China, Beijing. It was the National Time Service Center of CAS located in Lintong District, Xi'an, Shaanxi Province that originated and announced the standard time of China.

Figure A1. A stimulus as an example of Study 1.

Note. The stimulus was adapted based on an article by CCTV News Center (2021).

Your Blood Type May Be Changed Through Medical Intervention

As a rule, a person's blood type will be the same throughout life. However, with respect to certain cases, the blood type may be changed. For example, having a specific medical intervention between people with different blood types, the recipient's blood type changes to that of the provider, and the change is permanent.

Vague News

---

* All the stimuli were originally in Chinese, and here we translate the text into English. The complete material is available on request from the authors.

Your Blood Type May Be Changed Through Stem Cell Transplantation

As a rule, a person's blood type will be the same throughout life. However, with respect to certain cases, the blood type may be changed. For example, having a hematopoietic stem cell transplantation between people with different blood types, the recipient's blood type changes to that of the provider, and the change is permanent.

True News

Your Blood Type May Be Changed Through Heart Transplantation

As a rule, a person's blood type will be the same throughout life. However, with respect to certain cases, the blood type may be changed. For example, having a heart transplantation between people with different blood types, the recipient's blood type changes to that of the provider, and the change is permanent.

False News

Figure A2. A set of stimuli as an example of Study 2.

Note. The stimuli were created based on an article by DXY (2016).

Appendix B: The Grouping Results

Table B1. Number of Individuals in Each Group of Study 1.

			Experimental Group		
News Item	Background Knowledge	Tool Use	Control	More-Questions	Read-Aloud
Beijing Time	Without	Without	84	46	77
		With	17	14	5
	With	Without	29	23	34
		With	11	17	13
Fingerprints	Without	Without	111	50	80
		With	4	11	7
	With	Without	20	22	27
		With	6	17	15
Blizzard	Without	Without	96	61	75
		With	6	4	7
	With	Without	29	15	32
		With	10	20	15
PM2.5	Without	Without	107	62	96
		With	4	6	8
	With	Without	20	11	12
		With	10	21	13

Table B2. Number of Individuals in Each Group of Study 2.

			Cue Type	
Set	Background Knowledge	Tool Use	True	False
1	Without	Without	108	103
		With	2	3
	With	Without	29	30
		With	6	5
2	Without	Without	105	110
		With	6	2
	With	Without	31	40
		With	4	3
3	Without	Without	99	96
		With	1	5
	With	Without	24	19
		With	6	6